CHAPTER 24

Reading 24.1 **Ellis, A. (1996). Recognising and understanding spoken words. In *Human Cognitive Neuropsychology: a textbook with readings*. Hove: Psychology Press, chapter 6 (Extract pp. 143–6).**

Reading 24.2 **Fodor, J.A. (1991). Propositional attitudes. In *The Nature of Mind*. (ed. D. Rosenthal). Oxford: Oxford University Press, pp. 325–338 (Extract pp. 325–6).**

Reading 24.3 **Fodor, J.A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press (Extract pp. 97–9).**

Reading 24.4 **Fodor, J.A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press, (Extract pp. 106–8).**

Reading 24.5 **Millikan, R. (1995). Biosemantics: explanation in biopsychology. In *Philosophy of Psychology* (ed. C. Macdonald and G. Macdonald). Oxford: Blackwell, pp. 253–276 (Extracts pp. 253, 255–8).**

Reading 24.1 EXERCISE 1

From: Ellis, A. (1996). Recognising and understanding spoken words. In *Human Cognitive Neuropsychology: a textbook with readings*. Hove: Psychology Press, chapter 6 (Extract pp.143–6).

Introduction

Spoken language travels from speaker to hearer as a sound wave. That sound wave is an extremely rich source of information. Without ever seeing a speaker we can often deduce correctly that person's sex, region of origin (from their accent), emotional state (e.g. whether they are happy, sad, or angry), approximate age, and so on. If the speaker is someone known to us we may be able to identify him or her as an individual from their voice and way of talking. There is, of course, linguistic information encoded in the speech wave too. This includes information about individual words, but in addition the syntactic boundaries of sentences or clauses are often signalled by pauses or changes in voice pitch, and even the transition from one general topic to another may be marked in a similar way (Ellis & Beattie, 1986).

We shall, however, principally be concerned here with recognising spoken words and extracting their meaning. Imagine the simple case of recognising a single word, clearly articulated and spoken in isolation. Unless that word is a homophone (like *their* and *there, one* and *won*) its sound pattern will be unique to it. To identify the word a listener will need to have stored in memory all the sound patterns of words he or she knows, and be able to compare the pattern just heard with these stored patterns to find the best match. What we are proposing is another word store or lexicon, but this time one involved in the recognition rather than the production of spoken words. We shall call it the *auditory input lexicon*.

There are currently two views prevalent on how the auditory input lexicon might work. One theory proposes that the listener first identifies phonemes (individual speech sounds) in the acoustic wave, and then identifies the word from its constituent phonemes. According to this view individual entries in the auditory input lexicon would be activated by a prior set of phoneme recognisers (e.g. Rumelhart & McClelland, 1981). The second theory, advocated by Klatt (1979) and Marcus (1981) among others, holds that the input to the auditory word recognition system is a low-level, relatively unsegmented description of the speech waveform. While acknowledging that either (or neither) of these theories may turn out to be correct in the long run, we shall tentatively adopt the first as our working hypothesis.

We propose, in Fig. 6.1, that the first stage of auditory word recognition performed by an early *auditory analysis system* attempts to identify phonemes in the speech wave. The results of this analysis are transmitted to the auditory input lexicon where a match is sought against the stored characteristics of known

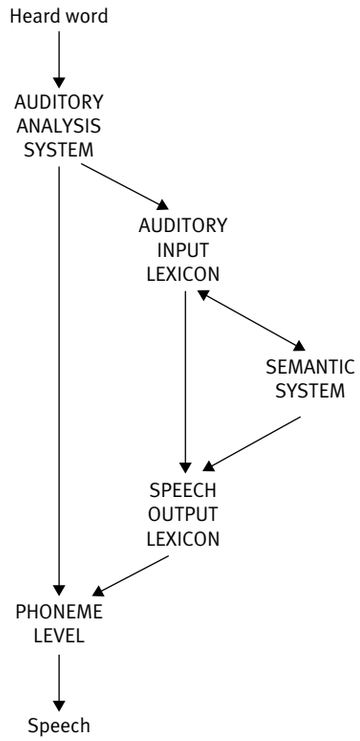

Fig 6.1. Simple functional model for the recognition, comprehension, and repetition of spoken words.

words. If the match is a good one, the appropriate recognition unit in the auditory input lexicon will be activated. It, in turn, will then activate the representation of the meaning of the heard word in the semantic system—the same semantic system that initiates the word production process in speaking via the same speech output lexicon and phoneme level that were discussed in the previous chapter. The arrow between the auditory input lexicon and the semantic system is bidirectional. This allows the semantic system to exert an influence upon the level of activity in the word-units which, in turn, provides a mechanism whereby the semantic context in which a word occurs can affect its ease of identification (see below).

One way to repeat a heard word would be to activate its entry in the speech output lexicon, release the phonemic form, and articulate it. This would be to take a route straight through Fig. 6.1. However, normal people can also repeat aloud unfamiliar words or non-words like "fep" or "flootil", for which there will be no entry in either the auditory input lexicon or the speech output lexicon. In Fig. 6.1, therefore, we need a by-pass route from the acoustic analysis system to the phoneme level. The by-pass route *must* be used to repeat unfamiliar words or non-words. It *could* be used for real words (treating them as if they were non-words), but real words can also be repeated via the input and output lexicons.

Figure 6.1 thus provides three "routes" between hearing a word and saying it. The first route is through word meanings and the

two lexicons; the second is provided by the direct link between the auditory analysis system and the phoneme level; and the third route is provided by the arrow linking the auditory input lexicon to the speech output lexicon. This would allow heard words to activate their entries in the speech output lexicons directly, without going via the representations of word meanings in the semantic system. We will admit here and now that the evidence for such a route is weak. It is included because it allows us to explain in Chapter 7 how certain patients are able to write words to dictation without understanding the meanings of those words.

McCarthy and Warrington (1984) employed a model similar to Fig. 6.1 to explain two different patterns of repetition performance observed in three aphasic patients they studied. Patient ORF was able to repeat words considerably better than non-words of the same length (85% correct for words *vs.* 39% for non-words). Non-words can only be repeated via the link between auditory analysis and the phoneme level, whereas words can be repeated via any of the routes through Fig. 6.1. ORF's superiority for words over non-words suggests some impairment to the connection between the auditory analysis system and the phoneme level, with the routes via the lexicons and semantics assisting word repetition. These latter routes were not completely intact, however: ORF had word-finding problems in speech and naming, and his word repetition was not perfect, being affected by both word frequency and length (with more common, shorter words being repeated better than less common, longer words). Errors in both word and non-word repetition were phonemic distortions. Examples of his errors in word repetition include "fosh" for "wash", "acroldut" for "acrobat", and "kwokutrain" for "porcupine".

Further evidence for a role of the lexical-semantic routes in ORF's repetition came from the observation that he was assisted in repeating a word if it followed an incomplete priming sentence (e.g. The monster was . . . *hideous*). A second patient (RAN) of McCarthy and Warrington (1984) was similar to ORF, but the third patient (ART) showed a different pattern. ART's spontaneous speech was "halting and effortful", with many phonemic distortions ("paraphasias"). His repetition of words was good, however, at around 90% correct, and was unaffected by frequency or length. Furthermore, he was actually *worse* at repeating words when they followed incomplete priming sentences than when the words were presented in isolation.

McCarthy and Warrington (1984) argue that ART's repetition is mediated by the connection between the auditory analysis system and the phoneme level. Although ART's non-word repetition was not assessed, one would expect on this interpretation no difference between word and non-word repetition (assuming that the by-pass route is insensitive to the wordness of phoneme sequences and that advantages for words over non-words come from contributions from the lexicons and semantics).

Figure 6.1 links speech input to speech output and thus incorporates a model for auditory-vocal repetition. Our central concern in this chapter will, however, be with the recognition and understanding of spoken words, rather than with their simple repetition. We shall consider a number of aphasias in which the recognition and understanding of spoken words is impaired, aphasias known in the literature as pure word deafness, word meaning deafness, auditory phonological agnosia, and deep dysphasia.

Reading 24.2 EXERCISE 2

From: Fodor, J.A. (1991). Propositional attitudes. In *The Nature of Mind* (ed. D. Rosenthal). Oxford: Oxford University Press, pp. 325–338 (Extract pp. 325–6).

Some philosophers (Dewey, for example, and maybe Austin) hold that philosophy is what you do to a problem until it's clear enough to solve it by doing science. Others (Ryle, for example, and maybe Wittgenstein) hold that if a philosophical problem succumbs to empirical methods, that shows it wasn't *really* philosophical to begin with. Either way, the facts seem clear enough: questions first mooted by philosophers are sometimes coopted by people who do experiments. This seems to be happening now to the question: "what are propositional attitudes?" and cognitive psychology is the science of note.

One way to elucidate this situation is to examine theories that cognitive psychologists endorse, with an eye to explicating the account of propositional attitudes that the theories presuppose. That was my strategy in Fodor (1975). In this paper, however, I'll take another tack. I want to outline a number of a priori conditions which, on my view, a theory of propositional attitudes (PAs) ought to meet. I'll argue that, considered together, these conditions pretty clearly demand a treatment of PAs as relations between organisms and internal representations; precisely the view that the psychologists have independently arrived at. I'll thus be arguing that we have good reasons to endorse the psychologists' theory even aside from the empirical exigencies that drove them to it. I take it that this convergence between what's plausible a priori and what's demanded ex post facto is itself a reason for believing that the theory is probably true.

Three preliminary remarks: first, I'm not taking 'a priori' all that seriously. Some of the points I'll be making are, I suppose, strictly conceptual, but others are merely self-evident. What I've got is a set of glaring facts about propositional attitudes. I don't doubt that we might rationally adopt an account of the attitudes which contravenes some, or maybe even all of them. But the independent evidence for such an account would have to be extremely persuasive or I, for one, would get the jitters. Second, practically everything I'll say about the attitudes has been said previously in the philosophical literature. All I've done is bring the stuff together. I do think, however, that the various constraints that I'll discuss illuminate each other; it is only when one attempts to satisfy them all at once that one sees how univocal their demands are. Finally, though I intend what I say to apply, mutatis mutandis, to PAs at large, I shall run the discussion pretty much exclusively on beliefs and wants. These seem to be the root cases for a systematic cognitive psychology; thus learning and perception are presumably to be treated as varieties of the fixation of belief, and the theory of action is presumably continuous with the theory of utility.

Here, then, are my conditions, with comments.

I. Propositional attitudes should be analyzed as relations. In particular, the verb in a sentence like 'John believes it's raining' expresses a relation between John and something else, and a token of that sentence is true iff John stands in the belief-relation to that thing. Equivalently, for these purposes, 'it's raining' is a term in 'John believes it's raining'. I have three arguments for imposing condition I, all of them inconclusive.

I-a) It's intuitively plausible. 'believes' looks like a two-place relation, and it would be nice if our theory of belief permitted us to save the appearances.

No doubt, appearances sometimes deceive. The 's' in 'Mary's sake' looks like expressing a relation (of possession) between Mary and a sake; but it doesn't, or so we're told. In fact, 'Mary's sake' doesn't look *very* relational, since *x's sake* would surely qualify as an idiom even if we had no ontological scruples to placate. There's something syntactically wrong with: Mary's sake is *Fer* than Bill's, 'Mary has a (little) sake', etc. For that matter, there's something syntactically wrong with 'a sake' *tout court*. Yet, we'd expect all such expressions to be well-formed if 'Mary's sake' contained a true possessive. 'Mary's sake' doesn't bear comparison with 'Mary's lamb'.

Still, there are some cases of *non*-idiomatic expressions which appear to be relational, but which, upon reflection, maybe aren't. 'Mary's voice' goes through the transformations even if 'Mary's sake' does not (Dennett, 1969). Yet there aren't, perhaps, such *things* as voices; and, if there aren't, 'Mary's voice' can't refer in virtue of a relation between Mary and one of them. I think it is fair to view the "surface" grammar as ontologically misleading in *these* cases, but only because we know how to translate into more parsimonious forms. 'Mary has a good voice (bad voice; little voice; better voice than Bill's)' goes over, pretty much without residue, into 'Mary sings well (badly, weakly, less well than Bill)'. If, however, we were *un*able to provide (or, anyhow, to envision providing) the relevant translations, what right would we have to view such expressions as ontologically promiscuous? 'Bill believes it's raining' is not an idiom, and there is, so far as anybody knows, no way of translating sentences nominally about beliefs into sentences of reduced ontological load. (Behaviorists used to think such translations might be forthcoming, but they were wrong.) We must, then, either take the apparent ontological commitments seriously or admit to playing fast and loose.

I-b) Existential Generalization applies to the syntactic objects of verbs of propositional attitude; from 'John believes it's raining' we can infer 'John believes something' and 'there is something that John believes' (viz., that it's raining). *EG* may not be *criterial* for ontological commitment, but it is surely a straw in the wind.

I-c) The only known alternative to the view that verbs of propositional attitude express relations is that they are (semantically) "fused" with their objects, and that view would seem to be hopeless.

Reading 24.3 EXERCISE 3

From: Fodor, J.A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press (Extract pp. 97–9).

Introduction

I suppose that sooner or later the physicists will complete the catalogue they've been compiling of the ultimate and irreducible properties of things. When they do, the likes of *spin, charm*, and *charge* will perhaps appear upon their list. But *aboutness* surely won't; intentionality simply doesn't go that deep. It's hard to see, in face of this consideration, how one can be a Realist about intentionality without also being, to some extent or other, a Reductionist. If the semantic and the intentional are real properties of things, it must be in virtue of their identity with (or maybe of their supervenience on?) properties that are themselves *neither* intentional *nor* semantic. If aboutness is real, it must be really something else.

And, indeed, the deepest motivation for intentional irrealism derives not from such relatively technical worries about individualism and holism as we've been considering, but rather from a certain ontological intuition: that there is no place for intentional categories in a physicalistic view of the world; that the intentional can't be *naturalized*. It is time that we should face this issue. What is it, then, for a physical system to have intentional states?

Let's, to begin with, try to get clear on just *where* the naturalization problem arises in the sort of account of propositional attitudes that I've been pushing. I've assumed that what most needs to be explained about the attitudes is that they have conditions of semantic evaluation; such facts as that beliefs have truth conditions, for example. Now, according to my story, you generate conditions for the semantic evaluation of an attitude by *fixing a context* for the tokenings of certain symbols; symbols which jointly constitute a system of mental representations. (The reader will recall that RTM is operative, and that RTM identifies token attitudes with relations between organisms and the token mental representations that they entertain.) So, then, what is it to fix a context for a system of mental representations?

Well, whatever else you have to do, you must at least specify an interpretation for items in the primitive nonlogical vocabulary of the language to which the symbols belong. For example, you fix a context for tokenings of the (Mentalese) expression 'this is water' by specifying—inter alia—that in the context in question the symbol 'water' expresses the property H_2O, or the property *XYZ*, or whatever. Granting an interpretation of the primitive nonlogical vocabulary, the business of generating conditions of evaluation for derived formulas can proceed by means which, though certainly not unproblematic, are at least familiar; viz., by the construction of a truth definition. In short: Given RTM, the intentionality of the attitudes reduces to the content of mental representations. Given a truth definition, the content of mental representations is determined by the interpretation of their primitive nonlogical vocabulary. So it's the interpretation of the primitive nonlogical vocabulary of Mentalese that's at the bottom of the pile according to the present view. Correspondingly, we would have largely solved the naturalization problem for a propositional-attitude psychology if we were able to say, in nonintentional and nonsemantic idiom, what it is for a primitive symbol of Mentalese to have a certain interpretation in a certain context.

Alas, I don't know how to carry out this program. But I see no principled reason why it can't be carried out; I even imagine that we might make a little progress within the foreseeable future. In particular, I think it's plausible that the interpretation of (primitive, nonlogical; from now on I'll omit these qualifiers) Mentalese symbols is determined by certain of their causal relations. For example, what makes it the case that (the Mentalese symbol) 'water' expresses the property H_2O is that tokens of that symbol stand in certain causal relations to water samples. Presumably if tokens of 'water' have a different interpretation on Twin-Earth (or, equivalently, if Twin-Earth counts as a *different context* for tokens of 'water'; or, equivalently, if tokens of 'water' are type-distinct from tokens of 'water2'; or, equivalently, if Mentalese2 counts as a different language from Mentalese), that is all because it's XYZ that bears to 'water2' tokens the sort of causal relations that H_2O bears to tokens of 'water.'

So the causal story goes. I think it points a promising route to the naturalization of such semantics as RTM requires. At a minimum, I think that some of the standard objections to that sort of story can be met; that's what I propose to argue in the following.

Here, then, are the ground rules. I want a *naturalized* theory of meaning; a theory that articulates, in nonsemantic and nonintentional terms, sufficient conditions for one bit of the world to *be about* (to express, represent, or be true of) another bit. I don't care—not just now at least—whether this theory holds for *all* symbols or for all things that represent. Maybe the occurrence of smoke expresses the proximity of fire; maybe the number of tree rings expresses the age of the tree; maybe the English predicate 'is red' expresses the property of being red; maybe the thermostat represents the temperature of the ambient air (see note 1). It's OK with me if any or all of this is so; but I'm not particularly anxious that the theory that naturalizes the semantic properties of mental representations should work for smoke, tree rings, or English words. On the contrary, I'm prepared that it should turn out that smoke and tree rings represent only relative to our interests in predicting fires and ascertaining the ages of trees, that thermostats represent only relative to our interest in keeping the room warm, and that English words represent only relative to our intention to use them to communicate our thoughts. I'm prepared, that is, that only mental states (hence, according to RTM, only mental representations) should turn out to have semantic properties *in the first instance*; hence, that a naturalized semantics should apply, strictu dictu, to mental representations only.

But it had better apply to them.

Reading 24.4 EXERCISE 5

From: Fodor, J.A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press, (Extract pp. 106–8).

How to Solve the Disjunction Problem

We need a way to break the symmetry between A-caused 'A' tokenings (which are, by hypothesis, true) and B-caused 'A' tokenings (which are, by hypothesis, false). In particular, we need a difference between A-caused 'A' tokenings and B-caused 'A' tokenings that can be expressed in terms of nonintentional and nonsemantic properties of causal relations; for nonintentional and nonsemantic properties of causal relations are all that the Crude Causal Theory of Content has to play with. My suggestion is that the teleological story was on the right track in appealing to the *counterfactual* properties of the causal relations between A's and 'A's, on the one hand, and B's and 'A's, on the other. Only the teleological story got hold of the wrong counterfactual properties.

It's an old observation—as old as Plato, I suppose—that falsehoods are *ontologically dependent* on truths in a way that truths are not ontologically dependent on falsehoods. The mechanisms that deliver falsehoods are somehow *parasitic on* the ones that deliver truths. In consequence, you can only have false beliefs about what you can have true beliefs about (whereas you can have true beliefs about anything that you can have beliefs about at all). So the intuition goes, and I think that there is something to it. What's more, I think that it points the way out of the disjunction problem.

Consider the following situation: I see a cow which, stupidly, I misidentify. I take it, say, to be a horse. So taking it causes me to effect the tokening of a symbol; viz., I say 'horse.' Here we have all the ingredients of the disjunction problem (set up, as it happens, for a token of English rather than a token of Mentalese; but none of the following turns on that). So, on the one hand, we want it to be that my utterance of 'horse' means *horse* in virtue of the causal relation between (some) 'horse' tokenings and horses; and, on the other hand, we *don't* want it to be that my utterance of 'horse' means *cow* in virtue of the causal relation between (some) 'horse' tokenings and cows. But if the causal relations are the same, and if causation makes representation, how can the semantic connections not be the same too? What we want is the situation in figure 1 (where the dashed line stands for the representation relation and the other lines stand for causal relations); but how are we to get what we want?

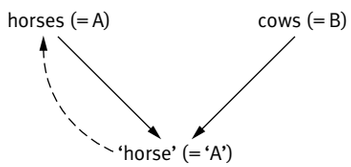

Figure 1

Answer: As previously remarked, the causal relations aren't identical in their counterfactual properties. In particular, misidentifying a cow as a horse wouldn't have led me to say 'horse' *except that there was independently a semantic relation between 'horse' tokenings and horses*. But for the fact that the word 'horse' expresses the property of *being a horse* (i.e., but for the fact that one calls *horses* 'horses', it would not have been *that* word that taking a cow to be a horse would have caused me to utter. Whereas, by contrast, since 'horse' does mean *horse*, the fact that horses cause me to say 'horse' does not depend upon there being a semantic—or, indeed, any—connection between 'horse' tokenings and cows.

From a semantic point of view, mistakes have to be *accidents*: if cows aren't in the extension of 'horse,' then cows being called horses can't be *required* for 'horse' to mean what it does. By contrast, however, if 'horse' didn't mean what it does, being mistaken for a horse wouldn't ever get a cow called 'horse.' Put the two together and we have it that the possibility of saying 'that's a horse' falsely presupposes the existence of a *semantic setup* for saying it truly, but not vice versa. Put it in terms of CCT, and we have it that the fact that cows cause one to say 'horse' depends on the fact that horses do; but the fact that horses cause one to say 'horse' does *not* depend on the fact that cows do.

So, the causal connection between cows and 'horse' tokenings is, as I shall say, *asymmetrically dependent* upon the causal connection between horses and 'horse' tokenings. So now we have a necessary condition for a B-caused 'A' token to be wild: B-caused 'A' tokenings are wild only if they are asymmetrically dependent upon non-B-caused 'A' tokenings.

What we've got so far is, in effect, a theory that understands wildness in terms of an empirical dependence among causal relations. Since all the notions employed are naturalistic, as per prior specifications, we could stop here. Alternatively, we can press the analysis further by reconstructing the notion of an empirical dependence in the familiar way, viz., by reference to subjunctives: If B-caused 'A' tokenings are wild—if they falsely represent B's as A's—then there *would be* a causal route from A's to 'A' even if there *were no* causal route from B's to 'A's; but there would be no causal route from B's to 'A's if there were no causal route from A's to 'A's.

Suppose that a counterfactual is true in a world iff its consequent is true in 'nearby' possible worlds in which its antecedent is true. (One possible world is 'near' another if, by and large, the laws that hold in the first also hold in the second. See Lewis, C.) So 'if I were smart I would be rich' is true here because I'm rich in the nearby possible worlds in which I'm smart. Then we can spell out the proposed condition on wild tokens as follows. In a world where B-caused 'A' tokens are wild (and express the property *A*), the nomic relations among properties have to be such that

1: A's cause 'A's

2: 'A' tokens are *not* caused by B's in nearby worlds in which A's *don't* cause 'A's.

3: A's cause 'A's in nearby worlds in which B's don't cause 'A's.

Reading 24.5 EXERCISE 6

From: Millikan, R. (1995). Biosemantics: explanation in biopsychology. In *Philosophy of Psychology* (ed. C. Macdonald and G. Macdonald). Oxford: Blackwell, pp. 253–276 (Extracts pp. 253, 255–8).

Extract 1: page 253

Causal or informational theories of the semantic content of mental states which have had an eye on the problem of false representations have characteristically begun with something like this intuition. There are some circumstances under which an inner representation has its represented as a necessary and/or sufficient cause or condition of production. That is how the content of the representation is fixed. False representations are to be explained as tokens that are produced under other circumstances. The challenge, then, is to tell what defines certain circumstances as the content-fixing ones.

Extract 2: pages 255–258

I fully agree, however, that an appeal to teleology, to function, is what is needed to fly a naturalist theory of content. Moreover, what makes a thing into an inner representation is, near enough, that its function is to represent. But, I shall argue, the way to unpack this insight is to focus on representation *consumption*, rather than representation production. It is the devices that *use* representations which determine these to be representations and, at the same time (*contra* Fodor), determine their content. If it really is the function of an inner representation to indicate its represented, clearly it is not just a natural sign, a sign that you or I looking on might interpret. It must be one that functions as a sign or representation *for the system itself*. What is it then for a system to use a representation *as* a representation?

The conception of function on which I shall rely was defined in my *Language, Thought, and Other Biological Categories* and defended in 'In defense of proper functions' under the label of 'proper function'. Proper functions are determined by the histories of the items possessing them; functions that were 'selected for' are paradigm cases. The notions 'function' and 'design' should not be read, however, as referring only to origin. Natural selection does not slack after the emergence of a structure but actively preserves it by acting against the later emergence of less fit structures. And structures can be preserved due to performance of new functions unrelated to the forces that originaly shaped them. Such functions are 'proper functions', too, and are 'performed in accordance with design'.

The notion 'design' should not be read—and this is very important—as a reference to innateness. A system may have been designed to be altered by its experience, perhaps to learn from its experience in a prescribed manner. Doing what it has learned to do in this manner is then 'behaving in accordance with design' or 'functioning properly'.

My term 'normal' should be read normatively, historically, and relative to specific function. In the first instance, 'normal' applies to explanations. A 'normal explanation' explains the performance of a particular function, telling how it was (typically) historically performed on those (perhaps rare) occasions when it was properly performed. Normal explanations do not tell, say, why it has been common for a function to be performed; they are not statistical explanations. They cover only past times of actual performance, showing how these performances were entailed by natural law, given certain conditions, coupled with the dispositions and structures of the relevant functional devices. In the second instance, 'normal' applies to conditions. A 'normal condition for performance of a function' is a condition, the presence of which must be mentioned in giving a full normal explanation for performance of that function. Other functions of the same organism or system may have other normal conditions. For example, normal conditions for discriminating tastes, and normal conditions for seeing very large objects are not the same as for seeing very small ones. It follows that 'normal conditions' must not be read as having anything to do with what is typical or average or even, in many cases, at all common. First, many functions are performed only rarely. For example, very few wild seeds land in conditions normal for their growth and development, and the protective colorings of caterpillars seldom actually succeed in preventing them from being eaten. Indeed, normal conditions might almost better be called 'historically optimal' conditions. (If normal conditions for proper functioning, hence survival and proliferation, were a statistical norm, imagine how many rabbits there would be in the world.) Secondly, many proper functions only need to be performed under rare conditions. Consider, for example, the vomiting reflex, the function of which is to prevent (further) toxification of the body. A normal condition for performance of this function is presence, specifically, of poison in the stomach, for (I am guessing) it is only under that condition that this reflex has historically had beneficial effects. But poison in the stomach certainly is not an average condition. (Nor, of course, is it a normal condition for other functions of the digestive system.)

If it is actually one of a system's functions to produce representations, as we have said, these representations must function as representations for the system itself. Let us view the system, then, as divided into two parts or two aspects, one of which produces representations for the other to consume. What we need to look at is the consumer part, at what it is to use a thing *as* a representation. Indeed, a good look at the consumer part of the system ought to be all that is needed to determine not only representational status but representational content. We argue this as follows. First, the part of the system which consumes representations must understand the representations proffered to it. Suppose, for example, that there were abundant 'natural information' (in Dretske's sense) contained in numerous natural signs all present in a certain state of a system. This information could still not

serve the system *as* information, unless the signs were understood by the system, and, furthermore, understood as bearers of whatever specific information they, in fact, do bear. (Contrast Fodor's notion that something could function like a representation without functioning like a representation of anything in particular.) So there must be something about the consumer that *constitutes* its taking the signs to indicate, say, *p, q* and *r* rather than *s, t* and *u*. But, if we know what constitutes the consumer's *taking* a sign to indicate *p*, what *q*, what *r*, etc., then, granted that the consumer's takings are in some way systematically derived from the structures of the signs so taken, we can construct a semantics for the consumer's language. Anything the signs may indicate *qua* natural signs or natural information carriers then drops out as entirely irrelevant; the representation-producing side of the system had better pay undivided attention to the language of its consumer. The sign producer's function will be to produce signs that are true *as the consumer reads the language*.

The problem for the naturalist bent on describing intentionality, then, does not concern representation production at all. Although a representation always is something that is produced by a system whose proper function is to make that representation correspond by rule to the world, what the rule of correspondence is, what gives definition to this function, is determined entirely by the representation's consumers.

For a system to use an inner item as a representation, I propose, is for the following two conditions to be met. First, unless the representation accords, *so* (by a certain rule), with a represented, the consumer's normal use of, or response to, the representation will not be able to fulfill all of the consumer's proper functions in so responding—not, at least, in accordance with a normal explanation. (Of course, it might still fulfill these functions by freak accident, but not in the historically normal way.) Putting this more formally, that the representation and the represented accord with one another, so, is a normal condition for proper functioning of the consumer device as it reacts to the representation. Note that the proposal is not that the content of the representation rests on the function of the representation or of the consumer, on what these do. The idea is not that there is such a thing as behaving like a representation of X or as being treated like a representation of X. The content hangs only on there being a certain condition that would be *normal* for performance of the consumer's functions—namely, that a certain correspondence relation hold between sign and world—whatever those functions may happen to be. For example, suppose the semantic rules for my belief representations are determined by the fact that belief tokens in me will aid the devices that use them to perform certain of their tasks in accordance with a normal explanation for success only under the condition that the forms or 'shapes' of these belief tokens correspond, in accordance with the said rules, to conditions in the world. Just what these user tasks are need not be mentioned.

Secondly, represented conditions are conditions that vary, depending on the *form* of the representation, in accordance with specifiable correspondence rules that give the semantics for the relevant *system* of representation. More precisely, representations always admit of significant transformations (in the mathematical sense), which accord with transformations of their corresponding representeds, thus displaying significant articulation into variant and invariant aspects. If an item considered as compounded of certain variant and invariant aspects can be said to be 'composed' of these, then we can also say that every representation is, as such, a member of a representational system having a 'compositional semantics'. For it is not that the represented condition is itself a normal condition for proper operation of the representation consumer. A certain correspondence between the representation and the world is what is normal. Coordinately, there is no such thing as a representation consumer that can understand only one representation. There are always other representations, composed other ways, saying other things, which it could have understood as well, in accordance with the same principles of operation. A couple of very elementary examples should make this clear.

First, consider beavers, who splash the water smartly with their tails to signal danger. This instinctive behavior has the function of causing other beavers to take cover. The splash means danger, because only when it corresponds to danger does the instinctive response to the splash on the part of the interpreter beavers, the consumers, serve a purpose. If there is no danger present, the interpreter beavers interrupt their activities uselessly. Hence, that the splash corresponds to danger is a normal condition for proper functioning of the interpreter beavers' instinctive reaction to the splash. (It does not follow, of course, that it is a usual condition. Beavers being skittish, most beaver splashes possibly occur in response to things not in fact endangering the beaver.) In the beaver splash semantic system, the time and place of the splash varies with, 'corresponds to', the time and place of danger. The representation is articulate: properly speaking, it is not a splash but a splash-at-a-time-and-a-place. Other representations in the same system, splashes at other times and places, indicate other danger locations.

Secondly, consider honey bees, which perform 'dances' to indicate the location of sources of nectar they have discovered. Variations in the tempo of the dance and in the angle of its long axis vary with the distance and direction of the nectar. The interpreter mechanisms in the watching bees—these are the representation consumers—will not perform their full proper functions of aiding the process of nectar collection in accordance with a normal explanation, unless the location of nectar corresponds correctly to the dance. So, the dances are representations of the location of nectar. The full representation here is a dance-at-a-time-in-a-place-at-a-tempo-with-an-orientation.